Mol Gen Genomics (2005) 273: 273–281
DOI 10.1007/s00438-005-1125-x

ORIGINAL PAPER

N. Atibalentja · S. Bekal · L. L. Domier · T. L. Niblack
G. R. Noel · K. N. Lambert

A genetic linkage map of the soybean cyst nematode *Heterodera glycines*

Received: 25 June 2004 / Accepted: 25 January 2005 / Published online: 6 April 2005

Abstract A genetic linkage map of the soybean cyst nematode (SCN) *Heterodera glycines* was constructed using a population of F_2 individuals obtained from matings between two highly inbred SCN lines, TN16 and TN20. The AFLP fingerprinting technique was used to genotype 63 F_2 progeny with two restriction enzyme combinations (*Eco*RI/*Mse*I and *Pst*I/*Taq*I) and 38 primer combinations. The same F_2 population was also genotyped for *Hg-cm-1* (*H. glycines* chorismate mutase-1), a putative virulence gene, using real-time quantitative PCR. Some of the markers were found to be distributed non-randomly. Even so, of the 230 markers analyzed, 131 could be mapped onto ten linkage groups at a minimum LOD of 3.0, for a total map distance of 539 cM. The *Hg-cm-1* locus mapped to linkage group III together with 16 other markers. The size of the *H. glycines* genome was estimated to be in the range of 630–743 cM, indicating that the current map represents 73–86% of the genome, with a marker density of one per 4.5 cM, and a physical/genetic distance ratio of between 124 kb/cM and 147 kb/cM. This genetic map will be of great assistance in mapping *H. glycines* markers to genes of interest, such as nematode virulence genes and genes that control aspects of nematode parasitism.

Keywords AFLP markers · Real-time PCR · Plant pathogen · Whole-genome amplification

Communicated by W.R. McCombie

N. Atibalentja · S. Bekal · L. L. Domier · T. L. Niblack
G. R. Noel · K. N. Lambert (✉)
Department of Crop Sciences,
University of Illinois at Urbana-Champaign,
1102 South Goodwin Avenue, Urbana, IL 61801, USA
E-mail: knlamber@uiuc.edu

L. L. Domier · G. R. Noel
USDA, ARS,
1102 South Goodwin Avenue, Urbana, IL 61801, USA

Introduction

The soybean cyst nematode (*Heterodera glycines*) is the most economically important pathogen of soybean (*Glycine max*) in the USA and in other soybean production areas around the world (Workneh et al. 1999; Wrather et al. 2001; Niblack et al. 2004). Annual yield losses from *H. glycines* have been estimated at over $2 billion, worldwide (Wrather et al. 2001). Soybean infection begins when a second-stage juvenile (J2) of *H. glycines* penetrates the root and migrates intracellularly to a region near the vascular cylinder, where it stimulates the formation of a complex feeding site, a syncytium (Endo 1992). Once the feeding site is established, the nematode feeds and molts to its adult stage. Males move out of the root in search of females to fertilize 10–15 days after invasion. Meanwhile, adult females, which remain attached to the plant root, enlarge as they develop, and eventually rupture the cortex and epidermis. At maturity, the female mates with one or more males and produces ~200–600 eggs. Typically, one-third of the eggs are laid and two-thirds remain inside the hardened body of the female forming the cyst. The *H. glycines* life cycle is completed when the J2 hatches from the egg and moves through the soil to infect a host root (Young 1992).

Current management practices to control *H. glycines* involve crop rotation with non-hosts and use of resistant cultivars. The inherent limitation of this strategy, however, is that soybean resistance to *H. glycines* is neither complete nor durable, as evidenced by the frequent occurrence of resistance-breaking populations of *H. glycines* (Koenning and Barker 1998). Disease development depends on the genetics of both the host and the pathogen. While significant progress has been made in the identification and mapping of *H. glycines* resistance genes in soybean (Weisemann et al. 1992; Concibido et al. 1994, 1997; Webb et al. 1995; Keim et al. 1997; Mudge et al. 1997), very little is understood about the genetics of virulence (ability to reproduce and cause

disease on resistant hosts) in *H. glycines*. Dong and Opperman (1997) crossed three *H. glycines* inbred lines that exhibited different levels of virulence on *H. glycines*–resistant soybean and found that virulence was controlled by independent single genes, which they called *ror* genes. During the last few years, a great deal of effort has been invested in the discovery of *H. glycines* parasitism genes (Gao et al. 2001, 2003; Wang et al. 2001; Bekal et al. 2003). The thousands of *H. glycines* ESTs that have been released will aid this endeavor (McCarter et al. 2003). Of the putative *H. glycines* parasitism genes that have been cloned, *H. glycines* chorismate mutase-1 (*Hg-cm-1*) is a good candidate for a virulence gene, since it displays polymorphisms that correlate with virulence of some *H. glycines* inbred lines on specific soybean indicator lines. *Hg-cm-1* is a member of a polymorphic gene family, and two different alleles, *Hg-cm-1*A and *Hg-cm-1*B, have been identified (Bekal et al. 2003). The *Hg-cm-1*A and *Hg-cm-1*B alleles segregated in a 1:2:1 genotypic ratio in F_2 populations derived from two highly inbred lines differing in virulence. When these F_2 populations were cultured on *H.glycines*–resistant soybean, *Hg-cm-1*A was preferentially selected, thus implicating it in *H.glycines* virulence (Lambert et al. 2005).

In cultivated plant species such as maize, potato, soybean and tomato, genetic linkage maps have become valuable tools in both fundamental and applied genetics (Tanksley et al. 1992; Shoemaker and Specht 1995; Vuylsteke et al. 1999). Genetic linkage maps, especially high-density maps, are particularly useful for map-based cloning of genes of biological and economic importance, such as virulence genes. In animal and plant breeding, high density maps have also been exploited for marker-assisted selection of desirable traits (Burr et al. 1983; Tanksley et al. 1989). Moreover, integrated genetic and physical maps have provided the necessary framework for whole-genome sequencing (The *Caenorhabditis elegans* Sequencing Consortium 1998; Marra et al. 1999; Waterston et al. 2002). The only genetic linkage map available for a plant-parasitic nematode is that of *Globodera rostochiensis* (Rouppe van der Voort et al. 1999).

Our objective was to construct a genetic linkage map for *H. glycines*, using the AFLP fingerprinting technique and an F_2 mapping population that originated from controlled matings between two *H. glycines* inbred lines. To show the usefulness of our map, we have determined the location of the *Hg-cm-1* locus on one of the linkage groups.

Materials and methods

Mating experiments

The F_2 mapping population used in this study originated from controlled matings between the *H. glycines* inbred lines TN16 and TN20, which have been maintained in the greenhouse for over 30 generations of selection by single cyst descent on the resistant soybean accessions PI209332 and PI437654, respectively. The techniques used for mating nematodes have been described elsewhere (Dong and Opperman 1997; Lambert et al. 2005). Briefly, virgin TN16 females and TN20 males were produced separately on the *H.glycines*–susceptible soybean cultivar Lee 74, grown in a hydroponic system. Plants infected with virgin TN16 females were transplanted individually into sterilized sandy soil to which TN20 males were added to mate. At maturity, gravid females were collected and crushed to release F_1 eggs. These eggs were bulked, and hatched J2 were used to infect Lee 74 plants on which the nematodes were allowed to mate randomly for one generation. The resultant gravid females were collected, and J2 worms hatched from F_2 eggs were used to infect Lee 74 plants, which were subsequently transferred to the hydroponic system. Virgin F_2 females were collected from these cultures and transferred individually into 1.5-ml microfuge tubes containing 10 µl of water, and stored at −80°C until used for DNA extraction.

DNA extraction and whole genome amplification

To extract genomic DNA from individual virgin F_2 females, the frozen nematodes were transferred onto dry ice, then briefly thawed and macerated with a tissue grinder (ISC BioExpress, Kaysville, UT, USA). The micropestle was rinsed with 400 µl of fresh Proteinase K extraction buffer [50 mM TRIS-HCl pH 8.0, 5 mM EDTA, 50 mM NaCl, 0.5% SDS, 1 mM DTT, 10 µg/ml tRNA, and 2 mg/ml Proteinase K (Invitrogen, Carlsbad, CA, USA)], and the mixture was incubated at 37°C for 1 h (Ausubel et al. 1992). The extract was transferred to 2-ml tubes containing Phase Lock Gel (Eppendorf, Hamburg, Germany), and 400 µl of phenol:chloroform:isoamyl alcohol (25:24:1) was added, followed by 5 min of centrifugation at high speed (16,000 *g*). The aqueous phase was then transferred to a fresh 1.5-ml microfuge tube, mixed with 1 µl of Pellet Paint NF (Novagen, Madison, WI, USA), and DNA was precipitated with 40 µl of 3 M sodium acetate (pH 5.2) and 800 µl of 100% ethanol for 2 min at room temperature. The mixture was centrifuged for 10 min at high speed, and the blue pellet was washed with 1 ml of 70% ethanol, followed by a 5-min centrifugation at high speed. Pellets were dried in a vacuum microcentrifuge at 65°C for 5 min, and stored at −20°C until used for whole genome amplification with the GenomiPhi DNA Amplification Kit (Amersham Biosciences, Sunnyvale, CA, USA). A similar extraction procedure was used to obtain genomic DNA from approximately 10,000 eggs each of the parent lines TN16 and TN20 (Bekal et al. 2003). For whole-genome amplification, each DNA pellet was resuspended in 1 µl of water and 9 µl of sample buffer before being subjected to the amplification reaction following the manufacturer's recommendations.

AFLP protocol

The AFLP reactions were conducted as described in Vos et al. (1995), with minor modifications. Two different enzyme combinations, *Eco*RI/ *Mse*I and *Pst*I/*Taq*I, were used. The sequences of the adapters and primers for *Eco*RI (E) and *Mse*I (M) were from Vos et al. (1995). The sequences for *Pst*I (P) and *Taq*I (T) were reported by Waugh et al. (1997) and Habu et al. (1997), respectively. For each enzyme combination, genomic DNA (~250 ng per reaction) from the parents TN16 and TN20, and GenomiPhi-amplified DNA from parents and F_2 progeny, were digested simultaneously with both endonucleases, and ligated with the corresponding adapters. The *Eco*RI/*Mse*I- and *Pst*I/*Taq*I-ligated DNA [1:5 dilution in TE buffer (10 mM TRIS-HCl, pH 8.0, 1 mM EDTA)] was pre-amplified with the primer combinations E-0/M-C and P-0/T-A, respectively. The pre-amplified DNA (1:20 dilution in TE buffer) was then used as the template in amplification reactions with primers that contained three selective nucleotides each. The *Eco*RI and *Pst*I primers were labeled with [$\gamma-^{33}$P]ATP. The products of selective amplification were fractionated in 5% denaturing polyacrylamide gels at constant power (50 W) for 4 h, and gels were exposed to Kodak BioMax MR films (Eastman Kodak, Rochester, NY, USA) for at least 3 days. For consistency, each gel was loaded first with the 30–330 bp AFLP DNA size standard (Invitrogen), followed by the two parental genomic DNAs, then the F_2 progeny DNAs, and finally the parental GenomiPhi-amplified DNAs.

Scoring of AFLP markers

Polymorphic AFLP fragments were scored in F_2 progeny as dominant markers, based on the presence or absence of bands on autoradiographs, following the A, B, C, D coding system for an F_2 intercross (Van Ooijen and Vooripps 2001). Only markers that were polymorphic between the parents, based on genomic DNA, and segregated among the F_2 progeny were considered. Marker names (e.g., E-AAG/M-CAA-142B) were formed such that the first and fifth letters indicate the enzyme and primer combinations used, each followed by the three selective nucleotides. The number in each marker name represents the approximate size (in bp) of the AFLP fragment, and the last letter indicates whether the marker was present in TN16 (A) or TN20 (B). The sizes of markers larger than 330 bp were estimated from scanned autoradiographs by interpolation using the GeneProfiler 4.0 software (Scanalytics, Fairfax, VA).

Mapping of the *Hg-cm-1* gene

Real-time quantitative PCR (Q-PCR) was used to determine the *Hg-cm-1* genotype of each of the F_2 progeny that were fingerprinted by AFLP. Genomic DNA from parents and GenomiPhi-amplified DNA from each F_2 individual were diluted to approximately 1 ng/μl before they were used for Q-PCR. The diluted DNA (5 μl per reaction) was transferred to a 96-well Q-PCR plate, and the assay was performed as previously described (Lambert et al. 2005), except that the allelic discrimination mode also was used to determine the *Hg-cm-1* genotype of F_2 individuals.

Linkage analysis

Linkage analysis was performed with the JoinMap 3.0 software (Van Ooijen and Vooripps 2001) with default options. All markers were analyzed for linkage, and recombination fractions were converted into map distances (centimorgans, cM) with the Kosambi function. A threshold LOD score of 3.0 was chosen for the construction of maps. Segregation ratio distortion was evaluated with a χ^2 test for goodness of fit based on the Mendelian expectations of 3:1 for AFLP markers and 1:2:1 for *Hg-cm-1*.

Genome size estimation, map coverage, and marker distribution along the map

Three different methods were used to estimate the length of the *H. glycines* genome (G) from linkage data. First, genome length was estimated as $G=[N(N-1)/K]X$, where N is the number of markers scored, X is the distance between two markers for which the value of the expected LOD is equal to a specified threshold Z (in this study, $Z=3.0$), and K is the number of marker pairs linked at the expected LOD Z or greater (Chakravarti et al. 1991, method 2; Hulbert et al. 1988). Second, genome length was determined by summing the lengths of all linkage groups, after each length had been adjusted by the factor $m+1/m-1$, where m is the number of framework markers on the linkage group (Chakravarti et al. 1991, method 4). Finally, map density (d), i.e., the average spacing between framework markers, was calculated by dividing the summed lengths of all linkage groups by the total number of intervals, which is equal to the number of framework markers minus the number of linkage groups. Twice the value of d was added to the length of each linkage group to account for chromosome ends beyond the terminal markers, and G was obtained by summing up the resulting lengths of all linkage groups (Fishman et al. 2001). For each of the three estimates of genome length, the expected map coverage c, i.e., the proportion of the genome within a distance d of a marker was calculated as $c=1-e^{-2dN/G}$ (Lange and Boehnke 1982). The actual map coverage was obtained by dividing the length of the framework map by the estimated genome length. Since the above estimation procedures assume a random distribution of markers along the map, a χ^2 test for goodness of fit was

performed to test this hypothesis (Rouppe van der Voort et al. 1997).

Results

AFLP fingerprinting

An F_2 mapping population in which the *Hg-cm-1*A and *Hg-cm-1*B alleles were segregating in a 1:2:1 genotypic ratio was obtained from controlled matings between two inbred lines of *H. glycines*, TN16 and TN20 (Lambert et al. 2005). Genomic DNA was extracted from individual virgin F_2 females and subjected to whole-genome-amplification to generate enough DNA for this study and for future mapping of *H. glycines* genes. The whole genome-amplification reaction yielded between 3 μg and 10 μg of DNA per individual, which provided sufficient DNA for multiple AFLP analyses. Two enzyme combinations, *Eco*RI/*Mse*I (E/M) and *Pst*I/*Taq*I (P/T), and 38 primer combinations (19 per enzyme combination) were used to genotype 63 F_2 progeny. For most of the primer combinations tested, nearly all of the bands (99% or higher) obtained with the parental genomic and whole genome-amplified DNA were shared, indicating that the latter DNA was suitable for AFLP analysis.

Considering all the 38 primer combinations, 1,577 (25%) of the 6,327 AFLP fragments generated from the inbred lines TN16 and TN20 (3,194 and 3,133 fragments, respectively) were polymorphic. Of these polymorphic fragments, 819 were TN16-specific and 758 were TN20-specific. The number of DNA fragments generated from each inbred line varied with the enzyme and primer combinations. With the E/M enzyme combination, the number of fragments amplified from TN16 ranged from 36 (E-ACG/M-CTA) to 119 (E-ACT/M-CAA); for TN20, the number of fragments varied from 39 (E-ACG/M-CTA) to 124 (E-AGC/M-CAC). The polymorphism index ranged from 17% (E-ACC/M-CTA) to 52% (E-AAG/M-CAT). With the P/T enzyme combination, the minimum and maximum numbers of fragments amplified from TN16 were 51 (P-AAG/T-AAA) and 154 (P-ACT/T-AAA), respectively, compared to 50 and 139 for TN20 with the same primer combinations. The polymorphism index here varied from 14% (P-AGC/T-ATT) to 37% (P-ACC/T-ATG).

Although 1,577 AFLP fragments were polymorphic between the parents TN16 and TN20, only 229 polymorphic markers could be confidently scored in 63 of the F_2 individuals analyzed. The 19 E/M primer combinations produced 111 (48.5%) of the 229 markers, with a range of two to ten markers and an average of six markers per primer combination. More than half (64) of the E/M markers came from TN20, compared with 47 from TN16. Two E/M primer combinations, E-AAG/M-CAT and E-AGG/M-CTT, produced only TN20 markers (two and five markers, respectively). The 19 P/T primer combinations yielded 118 (51.5%) markers, with a range of 1–13 markers and an average of six markers per primer combination. More P/T markers (63) originated from TN16 than from TN20 (55). The P-ACA/T-AAC primer combination produced no TN16 marker.

Segregation of AFLP markers

Nearly half (109) of the 229 AFLP markers segregated according to Mendelian expectations ($P > 0.05$). Of the 120 markers that showed skewed segregation, 45 (20% of the total) deviated from the expected ratios at $0.01 < P < 0.05$, 21 (9%) did so at $0.001 < P < 0.01$), and 54 (24%) deviated at $P < 0.001$. No trend was observed in the distribution of markers with distorted segregation between the enzymes or among the primer combinations.

Hg-cm-1 genotyping

To demonstrate the usefulness of the procedures described here for future mapping of *H. glycines* genes, we examined the segregation of the *Hg-cm-1*A/*Hg-cm-1*B allele pair in the mapping population. Of the 63 F_2 progeny genotyped, 23 were *Hg-cm-1*A/*Hg-cm-1*A, 29 were *Hg-cm-1*A/*Hg-cm-1*B, and 11 were *Hg-cm-1*B/*Hg-cm-1*B. These genotypic frequencies fit the 1:2:1 distribution, as shown by a χ^2 test for goodness of fit ($P > 0.05$). Consequently, the *Hg-cm-1* locus was treated as a co-dominant marker during linkage analysis.

Linkage analysis and map construction

All 230 markers (229 AFLPs and the *Hg-cm-1* locus) were analyzed for linkage, since the algorithm JoinMap used to calculate the LOD score for the recombination fraction is not affected by segregation ratio distortion (Van Ooijen and Voorips 2001). Of the 230 markers, 131 (57%) were mapped to ten linkage groups at a minimum LOD of 3.0 (Fig. 1). The *Hg-cm-1* locus mapped to linkage group III, along with 16 other markers. The number of markers per linkage group ranged from 3 to 39. Group length varied from 29 cM to 82 cM, resulting in a total map distance of 539 cM. The average marker spacing (map density) was 4.5 cM. Nine other markers were linked to at least one of the markers on the ten linkage groups at a minimum LOD of 3.0. However, due to low linkage information, the exact position of those markers could not be unambiguously determined. Therefore, they are not shown on the map. At a LOD threshold of 3.0, nine pairs involving 23 markers (some pairs resulted from triplets in which one of the markers could not be mapped due to low linkage information) were identified. Lowering the LOD threshold to 2.0 yielded 42 markers that were linked either to one of the framework markers or to a marker in one of the nine pairs. Two additional pairs of markers also were identified at a LOD threshold of 2.0. Finally,

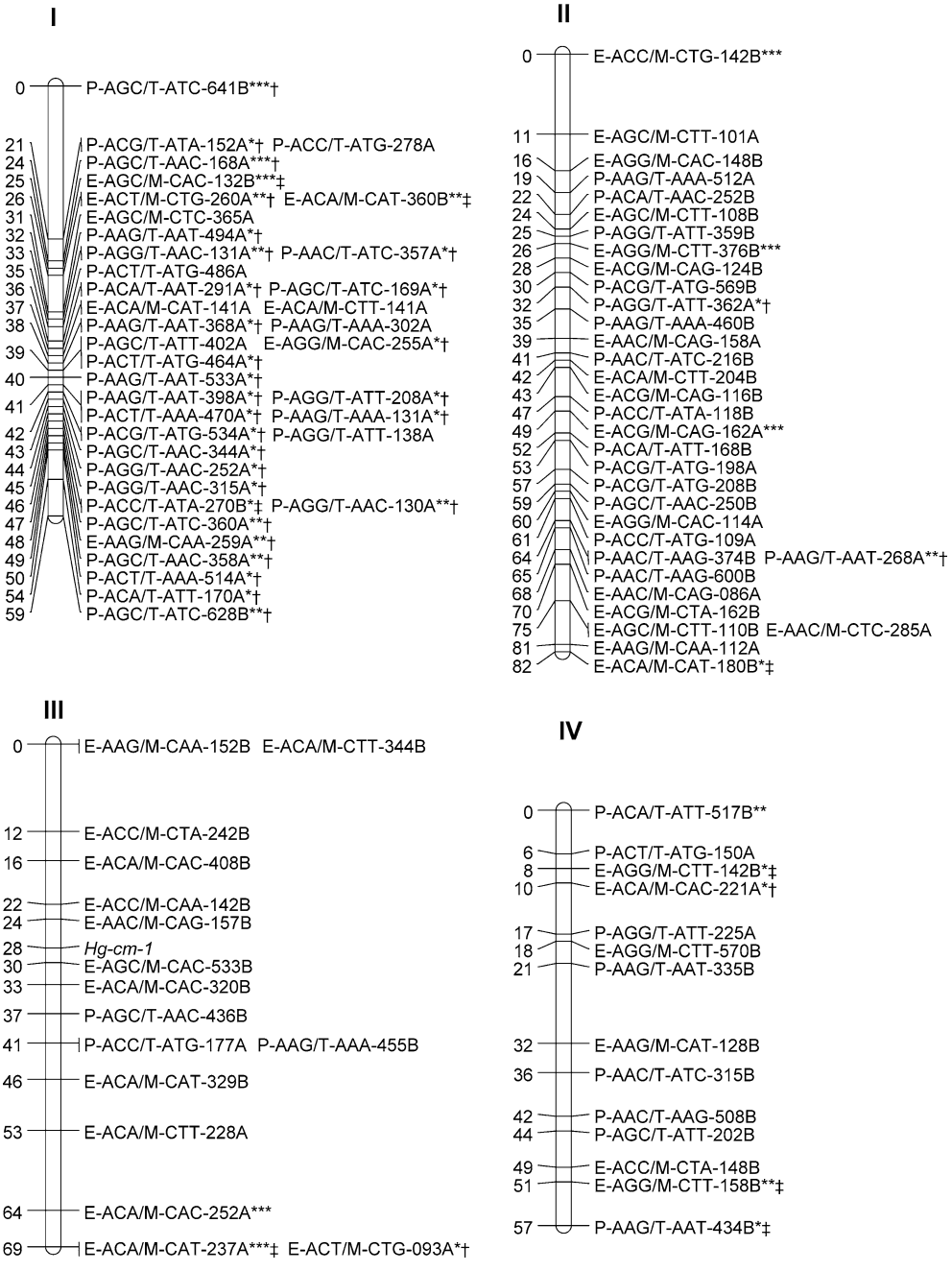


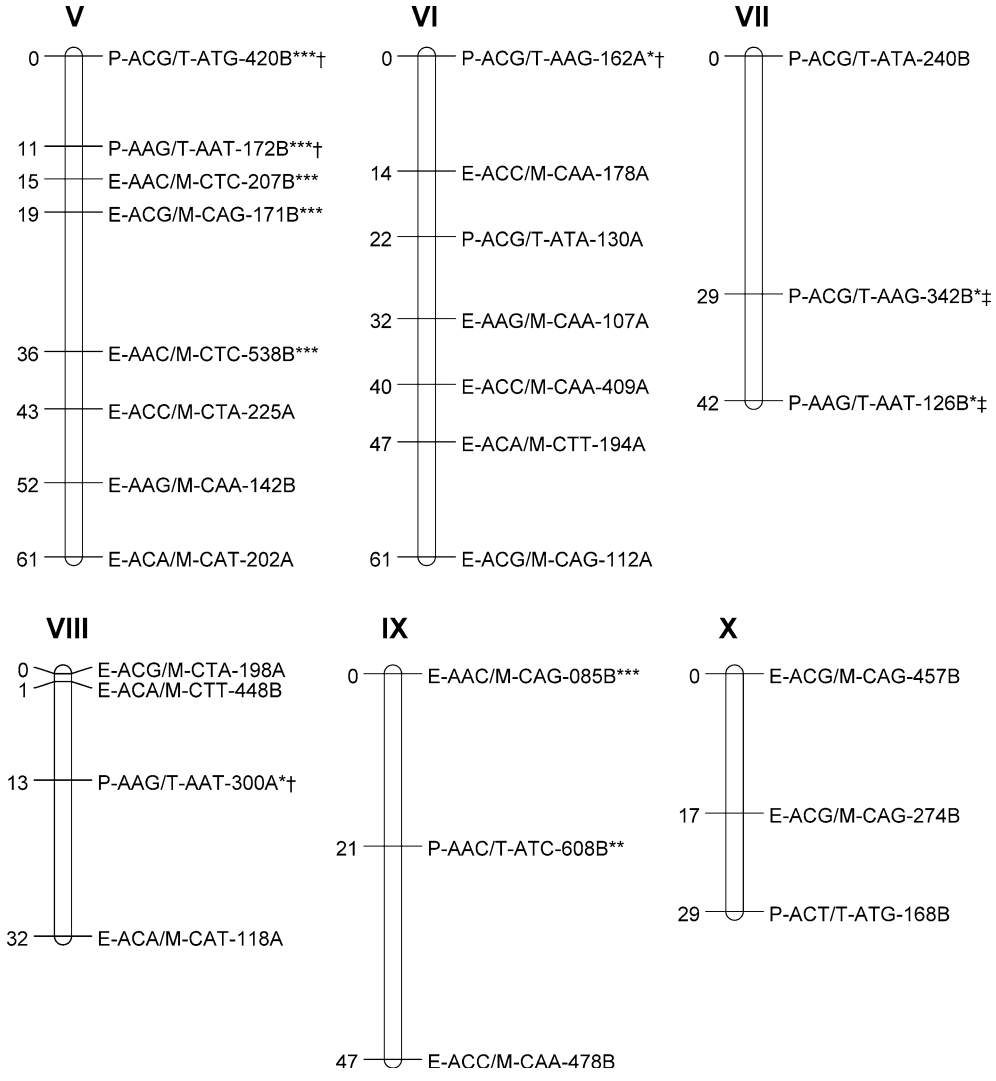
Fig. 1 A genetic linkage map of the soybean cyst nematode *Heterodera glycines*. The first and fifth letters of each AFLP marker name refer to the enzyme and primer combinations used (*E*, *M*, *P*, and *T* for *Eco*RI, *Mse*I, *Pst*I, and *Taq*I, respectively), which are followed by the three selective nucleotides used. The number indicates the approximate size (in bp) of the AFLP fragment, and the last letter indicates whether the marker was present in TN16 (**A**) or TN20 (**B**). *Hg-cm-1* is the chorismate mutase-1 gene. All markers were mapped at a minimum LOD of 3.0. For markers whose segregation patterns deviated from Mendelian expectations, the significance of the deviation is indicated by *asterisks* as follows: *, deviation at $P = 0.01$–0.05; **, deviation at $P = 0.001$–0.01; ***, deviation at $P < 0.001$. The symbols † and ‡ indicate deviation toward TN16 and TN20, respectively. *Asterisks* not followed by a symbol indicate a 1:1 segregation ratio. With the exception of E-ACT/M-CAA, all of the primer combinations tested are represented on the map

21 markers (9%) remained unlinked, 90% of which exhibited distorted segregation ratios. Two-thirds of the unlinked markers were TN20-specific, and 57% originated from the P/T enzyme combination.

Distribution of markers along the map

Overall, 62 (48%) of the 130 AFLP markers on the framework map originated from the E/M enzyme combination, and 68 (52%) markers came from the P/T enzyme combination. The distribution of markers among linkage groups varied markedly with the enzyme combination. Linkage groups III, V, VI, and VIII–X consisted predominantly (67–76%) of markers obtained with the E/M enzyme combination, whereas linkage groups I and VII mainly contained markers generated with the P/T enzyme combination (79% and 100%, respectively). Linkage groups II and IV had markers obtained with both enzyme combinations in nearly equal proportions. In general, each parent contributed half of the framework markers. However, the distribution of markers among linkage groups varied substantially depending on the parent from which the markers originated. Thus, 87, 100, and 75% of markers on linkage groups I, VI, and VIII, respectively, originated from TN16, while the other groups consisted mainly (64–100%) of TN20-specific markers. It is also obvious that

Fig. 1 (Contd.)



the majority of markers on linkage group I exhibited distorted segregation ratios, more so toward a deficit of TN20 null alleles (Fig. 1). Interestingly, linkage group I also happened to be the only group on which a significantly ($P < 0.05$) non-random distribution of markers was observed. Closer examination of that group revealed that clustering of markers occurred within the genomic region extending from marker P-AAG/T-AAT-494A to marker P-AGC/T-AAC-358A (Fig. 1). In this 17-cM segment, the number of markers (28) was significantly higher ($P < 0.05$) than expected under the hypothesis of random distribution.

Genome size estimation and map coverage

Taking all 230 markers scored on 63 F_2 progeny, 3,786 marker pairs were linked within 30 cM at a minimum expected LOD of 3.0. These figures correspond to a genome size estimate of 659 cM (Chakravarti et al. 1991, method 2; Hulbert et al. 1988). Based on the above estimate, the current map represents 82% of the *H. glycines* genome, which is close to the expected genome coverage of 96% with a map density of 4.5 cM. Adjusting the length of each linkage group according to method 4 of Chakravarti et al. (1991) resulted in an estimate of 743 cM and a genome coverage of 73%, compared to an expected genome coverage of 94%. In contrast, adjusting the length of each linkage group according to Fishman et al. (2001) gave an estimate of 630 cM and a genome coverage of 86% against an expectation of 96%.

Discussion

This is the first report of a genetic linkage map for *H. glycines* and only the second map for a plant-parasitic nematode. For the construction of the genetic linkage map of the potato cyst nematode, *G. rostochiensis*, Rouppe van der Voort et al. (1999) crossed individuals from two different field populations of the nematode. Due to the non-inbred nature of the cross, a "pseudo-F_2" mapping strategy was devised, in which only AFLP

markers from one parent were used. The availability of inbred lines of *H. glycines* allowed for the generation of a normal F_2 population and, therefore, offered the advantage that AFLP markers from both parents could be used for mapping. In both *H. glycines* and *G. rostochiensis*, mapping was based on segregation analysis of a series of sibling populations, each consisting of the progeny of a single generation of random mating (Rouppe van der Voort et al. 1999). The main pitfall associated with this method is that the DNA extracted from individual females is sufficient for only one AFLP mapping experiment involving one or two enzyme combinations, after which there is no possibility to map other loci or to investigate other traits at different times and in different environments. To circumvent this problem, we amplified the whole genome of each individual in our F_2 mapping population, obtaining 3–10 μg of DNA per cyst. With that amount of DNA, we were not only able to evaluate two different enzyme combinations, but also to map the *Hg-cm-1* locus and thus to demonstrate the usefulness of this population for future mapping of other *H. glycines* genes.

In its present version, the *H. glycines* linkage map comprises 131 markers distributed among ten linkage groups, for a total map distance of 539 cM. The size of the *H. glycines* genome appears to be between 630 cM and 743 cM, which is close to the 650 cM found for *G. rostochiensis* (Rouppe van der Voort et al. 1999). Thus, the map covers 73–86% of the genome, at a map density of 4.5 cM. Based on an estimated physical size of 92.5×10^3 kb (Opperman and Bird 1998), the physical/genetic distance ratio for *H. glycines* would fall between 124 kb/cM and 147 kb/cM. This ratio is within the size range of DNA fragments that can be cloned in bacterial artificial chromosomes.

The foregoing discussion assumes that recombination events are randomly distributed along the genome (Liu 1998). However, unlike the closely related species G. *rostochiensis* (Rouppe van der Voort et al. 1999), *H. glycines* exhibited a non-random distribution of markers along the map, notably on linkage group I. Clustering of markers has also been observed in the map of the free-living nematode, *C. elegans* (Brenner 1974; Edgley and Riddle 1993) and in many plant genomes (Tanksley et al. 1992; Rouppe van der Voort et al. 1997; Alonso-Blanco et al. 1998; Vuylsteke et al. 1999; Saal and Wricke 2002). In plant genomes, clustering of markers has been associated with suppression of recombination in centromeric regions. Unlike plant chromosomes, however, nematode chromosomes are holocentric and do not have a single defined spindle attachment region (Goldstein and Triantaphyllou 1980; Albertson and Thomson 1982, 1993; Rouppe van der Voort et al. 1996). Therefore, it seems less likely that the cluster of markers on linkage group I is associated with a centromere. The preponderance of P/T markers on linkage group I tends to suggest that this cluster might have resulted from the amplification of repetitive elements rich in *Pst*I and *Taq*I cleavage sites, such as transposons (Waugh et al. 1997). Several multicopy repeat elements including tandem repeats, inverted repeats, and transposons have been found scattered over nematode genomes (Barnes et al. 1995; McReynolds et al. 1986; The *C. elegans* Sequencing Consortium 1998; Williams et al. 2000). Linkage group I is also remarkable in that most of the markers on the group exhibited skewed segregation ratios. In *G. rostochiensis*, segregation distortion was attributed to differential viability of genotypes due to inbreeding depression (Rouppe van der Voort et al. 1999). It is also possible that certain genotypes generated in the matings were less fit as zygotes or individuals than others, which could have resulted in the under-representation of makers linked to those phenotypes. This hypothesis finds some support in the fact that neighboring markers on linkage group I and, to a lesser extent, on linkage groups V, VII, and IX are distorted in the same direction. The action of selection on less viable genotypes might also explain the striking deficit of TN20 alleles on linkage group I. If true, this would suggest that linkage group I mainly consists of markers associated with the robustness of TN16.

The fact that the linkage map reported here represents 73–86% of the *H. glycines* genome indicates that 14–27% of the genome remain to be covered with markers. The need for additional markers is apparent, since the current map contains one linkage group more than the haploid number of chromosomes reported for *H. glycines* (Lapp and Triantaphyllou 1972) and many pairs of markers are not linked to any of the markers on the map. That the *G. rostochiensis* map covered 65% of the genome with 66 AFLP markers (Rouppe van der Voort et al. 1999), whereas at most 86% of the *H. glycines* genome was covered with 131 markers may appear surprising at first glance. However, it becomes less so when one considers the way in which markers were distributed along the two maps. The distribution was random in *G. rostochiensis* but non-random in *H. glycines*. It is not surprising, therefore, that 30% of the 131 framework markers on the *H. glycines* map clustered in a single linkage group, leaving 91 markers for the remaining groups. Another possible explanation could be the difference in the information content of the mapping strategies used, i.e., backcross for *G. rostochiensis* versus F_2 progeny for *H. glycines*. With dominant markers such as those provided by AFLP, backcross models provide as much information as co-dominant markers, whereas F_2 models yield less linkage information (Knapp et al. 1995; Liu 1998).

In summary, the generation of the *H. glycines* genetic linkage map using whole genome-amplified genomic DNA from a *H. glycines* F_2 population has overcome a formidable obstacle that has hitherto prevented the genetic analysis of this important plant-parasitic nematode. This genetic map will be of great assistance in mapping *H. glycines* markers to genes of interest, such as nematode virulence genes and genes that control aspects of nematode parasitism.

Acknowledgements This study was supported by grants from the Illinois Soybean Program Operating Board (ISPOB), the North Central Soybean Research Program (NCSRP), the Soybean Disease Biotechnology Center (SDBC) at the University of Illinois at Urbana-Champaign, and the United Soybean Board (USB). Linkage analyses were conducted in the Computer Laboratory of Dr. T.R. Rocheford, Professor of Crop Sciences at the University of Illinois at Urbana-Champaign. We also wish to thank K. Colgrove, A. Colgrove, R. Heinz, and K. Hajsaid for their technical support. Mention of a trademark or proprietary product does not constitute a guarantee or warranty of the product by the USDA and does not imply its approval to the exclusion of other products that may also be suitable. This work has been carried out in compliance with the current laws governing genetic experimentation in the USA

References

Albertson DG, Thomson JN (1982) The kinetochores of *Caenorhabditis elegans*. Chromosoma 86:409–428

Albertson DG, Thomson JN (1993) Segregation of holocentric chromosomes at meiosis in the nematode, *Caenorhabditis elegans*. Chromosome Res 1:15–26

Alonso-Blanco C, Peeters AJM, Koomneef M, Lister C, Dean C, Van den Bosch N, Pot J, Kuiper MTR (1998) Development of an AFLP based linkage map of L *er*, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of L *er*/Cvi recombinant inbred line population. Plant J 14:259–271

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (eds) (1992) Short protocols in molecular biology, 2nd edn. Wiley, New York

Barnes TM, Kohara Y, Coulson A, Hekimi S (1995) Meiotic recombination, noncoding DNA and genomic organization in *Caenorhabditis elegans*. Genetics 141:159–179

Bekal S, Niblack TL, Lambert KN (2003) A chorismate mutase from the soybean cyst nematode *Heterodera glycines* shows polymorphisms that correlate with virulence. Mol Plant Microbe Interact 16:439–446

Brenner S (1974) The genetics of *Caenorhabditis elegans*. Genetics 77:71–94

Burr B, Evola SV, Burr FA, Beckmann JS (1983) The application of restriction fragment length polymorphism to plant breeding. In: Setlow JK, Hollaender A (eds) Genetic engineering, vol 5. Plenum Press, New York, pp 45–59

Chakravarti A, Lasher LK, Reefer JE (1991) A maximum likelihood method for estimating genome length using genetic linkage data. Genetics 128:175–182

Concibido VC, Denny RL, Boutin SR, Hautea R, Orf JH, Young ND (1994) DNA marker analysis of loci underlying resistance to soybean cyst nematode (*Heterodera glycines* Ichinohe). Crop Sci 34:240–346

Concibido VC, Lange DA, Denny RL, Orf JH, Young ND (1997) Genome mapping of soybean cyst nematode resistance genes in 'Peking', PI90763, and PI88788 using DNA markers. Crop Sci 37:258–264

Dong K, Opperman CH (1997) Genetic analysis of parasitism in the soybean cyst nematode *Heterodera glycines*. Genetics 146:1311–1318

Edgley ML, Riddle DL (1993) The nematode *Caenorhabditis elegans*. In: O'Brien SJ (ed) Genetic maps: locus maps of complex genomes, 6th edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp 3.281–3.318

Endo BY (1992) Cellular responses to infection. In: Riggs RD, Wrather JA (eds) Biology and management of the soybean cyst nematode. APS Press, St. Paul, pp 37–49

Fishman L, Kelly AJ, Morgan E, Willis JH (2001) A genetic map in the *Mimulus guttatus* species complex reveals transmission ratio distortion due to heterospecific interactions. Genetics 159:1701–1716

Gao B, Allen R, Maier T, Davis EL, Baum TJ, Hussey RS (2001) Identification of putative parasitism genes expressed in the esophageal gland cells of the soybean cyst nematode *Heterodera glycines*. Mol Plant-Microbe Interact 14:1247–1254

Gao B, Allen R, Maier T, Davis EL, Baum TJ, Hussey RS (2003) The parasitome of the phytonematode *Heterodera glycines*. Mol Plant Microbe Interact 16:720–726

Goldstein P, Triantaphyllou AC (1980) The ultrastructure of sperm development in the plant-parasitic nematode *Meloidogyne hapla*. J Ultrastr Res 71:143–153

Habu Y, Fuka-Tanaka S, Hisatomi Y, Iida S (1997) Amplified restriction fragment length polymorphism mRNA fingerprinting using a single restriction enzyme that recognizes a 4-bp sequence. Biochem Biophys Res Comm 234:516–521

Hulbert SH, Ilott TW, Legg EJ, Lincoln SE, Lander ES, Michelmore RW (1988) Genetic analysis of the fungus, *Bremia lactucae*, using restriction fragment length polymorphisms. Genetics 120:947–958

Keim P, Schupp JM, Travis SE, Clayton K, Zhu T, Shi L, Ferreira A, Webb DM (1997) A high density genetic map of soybean based upon AFLP markers. Crop Sci 37:537–543

Knapp SJ, Holloway JL, Bridges WC (1995) Mapping dominant markers using F_2 matings. Theor Appl Genet 91:74–81

Koenning SR, Barker KR (1998) Survey of *Heterodera glycines* races and other plant-parasitic nematodes on soybean in North Carolina. Suppl J Nematol 30:569–576

Lambert KN, Bekal S, Domier LL, Niblack TL, Noel GR, Smyth C (2005) Selection of *Heterodera glycines chorismate mutase-1* alleles on nematode-resistant soybean. Mol Plant Microbe Interact (in press)

Lange K, Boehnke M (1982) How many polymorphic genes will it take to span the human genome? Am J Hum Genet 34:842–845

Lapp NA, Triantaphyllou AC (1972) Relative DNA content and chromosomal relationships of some *Meloidogyne*,*Heterodera*, and *Meloidodera* spp. (Nematoda: Heteroderidae). J Nematol 4:287–291

Liu BH (1998) Statistical genomics: linkage, mapping, and QTL analysis. CRC Press, Boca Raton

Marra M et al (1999) A map for sequence analysis of the *Arabidopsis thaliana* genome. Nat Genet 22:265–270

McCarter JP, Mitrevia M, Clifton SW, Bird DM, Waterston RH (2003) Nematode gene sequences: update for December 2003. J Nematol 35:465–469

McReynolds LA, DeSimone SM, Williams SA (1986) Cloning and comparison of repeated DNA sequences from the human filarial parasite *Brugia malayi* and the animal parasite *Brugia pahangi*. Proc Natl Acad Sci USA 83:797–801

Mudge J, Gregan PB, Kenworthy JP, Orf JH, Young ND (1997) Two microsatellite markers that flank the major soybean cyst nematode resistance locus. Crop Sci 37:1611–1615

Niblack TL, Tylka GL, Riggs RD (2004) Nematode pathogens of soybean. In: Boerma HR, Specht JE (eds) Soybeans: improvement, production, and uses, 3rd edn. American Society of Agronomy, Madison, pp 821–851

Opperman CH, Bird DM (1998) The soybean cyst nematode, *Heterodera glycines*: a genetic model system for the study of plant-parasitic nematodes. Curr Opin Plant Biol 1:342–346

Rouppe van der Voort JNAM, van Enckevort ELJG, Pijnacker LP, Helder J, Gommers FJ, Bakker J (1996) Chromosome number of the potato cyst nematode *Globodera rostochiensis*. Fundam Appl Nematol 19:369–374

Rouppe van der Voort JNAM, van Zandvoort P, van Eck HJ, Folkertsma RT, Hutten RB, Draaistra J, Gommers FJ, Jacobsen E, Helder J, Bakker J (1997) Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. Mol Gen Genet 255:438–447

Rouppe van der Voort JNAM, van Eck HJ, van Zandvoort PM, Overmars H, Helder J, Bakker J (1999) Linkage analysis by genotyping of sibling populations: a genetic map for the potato cyst nematode constructed using "pseudo-F2" mapping strategy. Mol Gen Genet 261:1021–1031

Saal B, Wricke G (2002) Clustering of amplified fragment length polymorphism markers in a linkage map of rye. Plant Breed 121:117–123

Shoemaker RC, Specht JE (1995) Integration of the soybean molecular and classical genetic linkage group. Crop Sci 35:436–446

Tanksley SD et al (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics 132:1141–1160

Tanksley SD, Young ND, Paterson AH, Bonierbale MW (1989) RFLP mapping in plant breeding: new tools for an old science. Biotechnology 7:257–264

The *C. elegans* Sequencing Consortium (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. Science 282:2012–2018

Van Ooijen JW, Voorrips RE (2001) JoinMap 3.0, software for the calculation of genetic linkage maps. Plant Research International, Wageningen

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Vuylsteke M, Mank R, Antonise R, Bastiaans E, Senior ML, Stuber CW, Melchinger AE, Lübberstedt T, Xia XC, Stam P, Zabeau M (1999) Two high-density AFLP linkage maps of *Zea mays* L.: analysis of distribution of AFLP markers. Theor Appl Genet 99:921–935

Wang X, Allen R, Ding X, Goellner M, Maier T, de Boer J, Baum TJ, Hussey RS, Davis EL (2001) Signal peptide-selection of cDNA cloned directly from esophageal gland cells of the soybean cyst nematode *Heterodera glycines*. Mol Plant Microbe Interact 14:536–544

Waterston RH et al (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420:520–562

Waugh R, McLean K, Flavell AJ, Pearce SR, Kumar A, Thomas BBT, Powel W (1997) Genetic distribution of *Bare-1*-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphism (S-SAP). Mol Gen Genet 253:687–694

Webb DM, Baltazar BM, Rao-Arelli AP, Schupp J, Clayton K, Keim P, Beavis WD (1995) Genetic mapping of soybean cyst nematode race 3 resistance loci in the soybean PI437654. Theor Appl Genet 91:574–581

Weisemann JM, Mattews BF, Devine TE (1992) Molecular markers located proximal to the soybean cyst nematode resistance gene, *Rhg4*. Theor Appl Genet 85:136–138

Williams SA, Lizotte-Waniewski MR, Foster J, Guiliano D, Daub J, Scott AL, Slatko B, Blaxter ML (2000) The filarial genome project: analysis of the nuclear, mitochondrial and endosymbiont genomes of *Brugia malayi*. Int J Parasitol 30:411–419

Workneh F, Tylka GL, Yang XB, Faghihi J, Ferris JM (1999) Regional assessment of soybean brown stem rot, *Phytophthora sojae*, and *Heterodera glycines* using area-frame sampling: prevalence and effects of tillage. Phytopathology 89:204–211

Wrather JA, Anderson TR, Arsyad DM, Tan Y, Ploper LD, Porta-Puglia A, Ram HH, Yorinori JT (2001) Soybean disease loss estimates for the top ten soybean-producing countries in 1998. Can J Plant Pathol 23:115–121

Young LD (1992) Epiphytology and life cycle. In: Riggs RD, Wrather JA (eds) Biology and management of the soybean cyst nematode. APS Press, St. Paul, pp 27–36